大数据可视化分析在支撑智库研究中的应用与创新*

熊励　王锟　赵明辉
上海大学管理学院 上海 200444

摘要：[**目的/意义**]大数据环境下的智库研究主要针对海量的信息进行有效采集、处理、分析，挖掘信息中蕴含的逻辑规律，并且将大量的无规则数据可视化表示，这是当今智库研究的热点。[**方法/过程**]文章通过分析国内外智库和大数据可视化研究现状，整理了目前大数据可视化的智库研究方法的相关理论。分析大数据可视化方法在智库研究中的应用与创新，并运用 CiteSpace II 软件分析我国智库研究现状。[**结果/结论**]结果表明：我国目前对智库研究的方法还比较单一，大数据可视化研究涉及多学科和行业应用。大数据可视化的智库研究方法发展迅速，包括支持智库大数据可视化分析的海量数据处理、可视化算法、图形可视化、人机交互、分布式处理等技术。大数据可视化分析在智库研究中的应用有助于智库数据高效处理、决策科学准确化、复杂图形可视化、复杂问题简化、用户交互良好、资源集聚共享、提升智库作用和影响力。

关键词：智库　大数据　可视化　可视化分析　研究方法

分类号：TP3-05　C912

Application and Innovation of Big Data Visualization Analysis in Supporting Think Tanks Research

Xiong Li　Wang Kun　Zhao Minghui
School of Management, Shanghai University, Shanghai 200444

Abstract: [Purpose/significance] Think tanks research in the big data environment aims at collecting, processing, analyzing, and digging out the logical rules contained in the massive information effectively, and representing large amounts of irregular data by visualization methods, which is a hot topic of research on think tanks. **[Method/process]** This article started with analyzing the current research status of think tanks and the big data visualization at home and abroad, and related theories of think tanks research methods based on big data visualization were sorted out. The application and innovation of big data visualization methods in supporting think tanks research were analyzed, and CiteSpace II software was used to analyze the research status of think tanks research in China. **[Result/conclusion]** The results show that the research methods of think tanks in China were relatively simple at present, and big data visualization research has two features: multi-disciplinary and multi-industry. Think tanks research methods based on big data visualization have developed rapidly, including techniques of the massive data process, the visualization algorithm, the graphic visualization, the human-computer interaction, the distributed processing and so on, which can support the big data visualization analysis in think tanks research. The application of big data visualization analysis in think tanks research can help think tanks process data efficiently, make decisions scientifically and accurately, visualize complex graphics, simplify complex problems, enhance users' interaction well, agglomerate and share resources, and enhance their role and influence in the application research of think tanks.

Keywords: think tank　big data　visualization　visualization analysis　research methods

***本文系**上海高校智库内涵建设设计划项目（项目编号：N.62-0129-17-201）研究成果之一。

作者简介：熊励（ORCID：0000-0002-6527-0517），上海大学管理学院教授，博士生导师，Email：xiongli8@shu.edu.cn；王锟（ORCID：），上海大学管理学院博士研究生，Email:1549131984@qq.com；赵明辉（ORCID：），上海大学管理学院硕士研究生，Email：mail@zhaominghui.me。

1 引言

智库研究实际上是建立在科学数据分析之上并对未来研究形势及走向进行预判，具有竞争力和影响力。随着智库研究所面对的问题日益复杂，如何做到智库的专业化和细致化以应对新型智库的挑战，将会成为决定智库机构今后成败的关键 [1] 。早在 2015 年 1 月，中共中央办公厅、国务院办公厅印发的《关于加强中国特色新型智库建设的意见》中就强调了政策咨询的重要性，需要以改革创新为动力，努力建设具有中共特色的新型智库体系，从而大力推动中国未来特色新型智库研究[2] 。随着互联网的快速发展，大数据的收集、存储、分析为我国的智库研究带来了新的研究动力和挑战。尤其是大数据可视化方法的应用，从政策运行、政府管理与决策等视角，基于海量数据处理、分析的可视化方法，帮助决策者和公众理解关键问题，获取有效信息，这将是未来智库研究的基本趋势。

1.1 智库研究现状

目前，智库研究已经进入了快速发展阶段。在机构数量方面，2016年2月9日，宾夕法尼亚大学发布的《2015年全球智库报告》（2015 Global Go To Think Tank Index Report）指出：全球智库机构数量高达6846家，其中美国智库数量位居全球第一，为1835家，我国智库数量位居全球第二，为435家[3] 。

国内外学者均对智库做了深入研究，Jacques Peter 就曾针对美国传统智库的发展情况，采用科学计量方法统计、测算智库数据，分析美国传统智库的运行机制[4]。国内学者吕红基于 WOS 对国际智库研究态势进行文献计量分析和信息挖掘，总结出目前国际智库的研究热点：（1）智库集团的发展运行现状及其对社会的影响力；（2）智库的角色定位与决策作用；（3）智库分类、领域及功能研究；（4）智库案例分析[5]。

通过文献数量的研究，能在一定程度上表征出研究领域的变化情况。因此对国内外文献进行检索，以相关文献发文量进行对比，了解智库研究的趋势。国内文献检索选取中国知网（CNKI）的CSSCI期刊数据库，以篇名为“智库”精确检索期刊文章，删除会议通知、报告等信息，时间区间为2000~2016年，共721篇，国外文献检索选取WOS核心合集数据库，以“Think tanks”为主题，文章类型为“Article”进行检索，时间区间为2000~2016年，共384篇，检索时间为2017年7月19日，相关文献的发文量检索结果如图1所示。

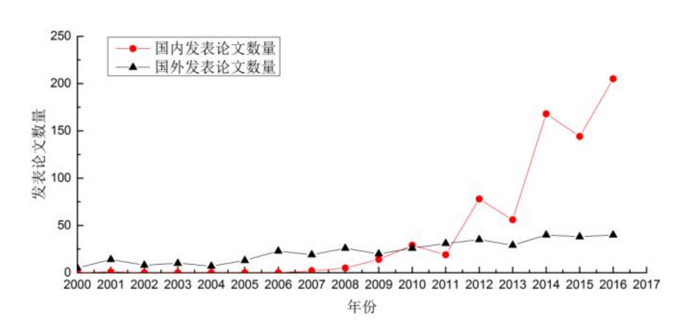


图 1 国内外智库研究发文量对比

Figure 1 The quantitative comparison of published articles in think tanks research at home and abroad

从图 1 可看出，2011 年之前国外对智库研究的文献数量相比国内较多，说明国外学者对智库研究较早，有一定的文献积累量，发表相关论文的数量在过去的 16 年间呈平稳增长趋势。而国内对智库的研究相对较晚，本世纪初才逐渐开始，但从 2011 年后相关论文数量逐渐增加。

1.2 大数据可视化研究现状

大数据可视化分析是目前大数据分析的重要方法之一，大规模的数据通过分析处理，并运用图形、图像方式直观展示数据信息。这些数据通常来自数据库、html、XML 文件、源代码等，通过存储并分析数据，实现数据的统计分析、知识表示、用户交互、视觉表现、人类认知、感知并探索决策[6-11]。大数据可视化技术充分利用数据分析，结合机器与人的优势分析问题、解决问题，实现视觉感知[12]。

大数据可视化方法来源于最初的科学可视化[13]和随着信息增长产生的信息可视化[14]，二者合称为数据可视化。Nightingale 曾运用传统数据可视化方法搜集、分析关于医疗和公共卫生的信息，其中运用最多的是饼图 [15]。随着互联网技术的发展和信息量的增加，大数据的可视化已经不再拘泥于传统的图形可视化，基于海量数据，逐渐在内容丰富的图形中展现交互功能。交互式数据可视化有助于复杂数据的选择和表示，能帮助用户和系统之间进行交互作用[16-17]。由于行业应用和数据类型的不同，大数据可视化分析方法也不同，其运用范围涉及政治、经济、金融、医疗等行业，处理的数据类型包括结构化数据、文本、网络、图像等数据 [18]。

1.3 问题提出

通过对国内外智库研究文献的对比分析，并进行研究现状的梳理，对比国内外智库研究的热点和趋势发现，国外智库研究起步较早，而我国智库研究虽有一定成果，但研究方法还比较单一，还未形成成熟的研究体系。大数据时代的到来为我国的智库研究提出了新的挑战，如何进行大数据驱动的智库研究？目前大数据可视化方法可针对海量数据直观展示信息，其应用涉及多种数据类型和行业决策分析，涉及多学科交叉研究，因此如何运用大数据可视化分析支撑智库研究的应用和创新？

文章通过对智库研究方法和理论进行阐述，并梳理大数据可视化的智库研究方法和理论，结合实际分析大数据可视化视角下的智库研究方法应用和创新，运用 CiteSpace II 软件对知网中部分智库研究文献进行可视化分析，并对结果进行展望。

2 智库研究方法发展动态

美国学者保罗·迪克森曾在其 1971 年出版的《Think Tanks》首次对智库进行了研究分析，并对智库的作用进行分析。随后越来越多的人意识到智库的重要性，国内外学者纷纷对智库进行研究，研究领域涵盖政府政策、经济、教育等诸多方面，智库的研究方法也在逐渐增多的研究中不断演变。

2.1 传统的计量统计方法

国外学者最初对智库研究主要是在理论研究的基础上，以定性分析法和定量分析法为基础进行的。Abelson 针对智库的影响力，运用定量分析的方法测度智库对社会、政治、经济的影响力 [19]。与此同时，问卷调查、专家访谈、访问等方法也逐渐运用到智库研究中，Rich 首先将统计学中的回归分析方法运用到智库研究中，对美国的医疗改革、政府政策制定过程中智库的作用进行了实证研究。[20]为了对国内外智库研究的态势进行分析，赵蓉英等运用文献的统计计量方法分析智库研究文献累积量，把握当今研究态势 [21]。从对智库理论研究到统计分析，传统的计量方法在智库研究中起到了重要作用，为大数据时代下的新兴的智库研究方法奠定了基础。

2.2 大数据环境下的智库研究方法

互联网和信息技术的快速发展使信息呈指数倍增长，智库信息也随之快速增长，为信息的收集、存储、分析带来挑战。潘教峰教授结合其对智库研究和实践经验，提出 DIIS 智库研究方法，认为智库研究只有走完 Data-Information-Intelligence-Solution 全过程才能真正提出有效服务决策的科学决策和咨询建议 [22]。

大数据环境下，数据驱动的思维逐渐在智库研究中扩散，大数据可视化分析方法得到灵活运用，实现海量数据分析和图形直观展示。如构建国际智库研究的国家地区间合作情况网络图谱，直观展示合作情况，并可绘制高频关键词的共现网络图 [5]，将可视化与文献分析结合进行国内外智库研究对比分析，把握当前智库研究的趋势和热点 [23]。

网络图谱能直观的展示智库研究的热点和趋势，为政府、教育机构等决策提供依据。目前对智库研究进行可视化分析的工具有 Pajek(合作网络图谱软件)、CiteSpace，VOSviewer 等。智库研究对政治、经济、教育的重要性正在逐渐增强，且大数据可视化的研究方法也较为方便直观，更加有利于决策者把握智库研究的热点和趋势。

2.3 大数据可视化的智库研究方法

大数据可视化技术是目前智库研究中进行数据表示、数据处理、决策分析的重要手段。目前大数据可视化的智库研究方法主要包括智库海量非结构化数据处理技术、智库研究大数据可视化算法技术、复杂智库信息的图形可视化技术、支持智库研究可视分析的人机交互技术、可视化的智库研究中分布式处理技术等。

智库海量非结构化数据处理技术是分析智库信息、数据的重要基础，主要针对海量非结构化数据的量化分析，目前大数据分析平台可以实现对实时采集到的海量非结构化数据的挖掘分析[24]。同时云存储技术的快速发展可提高对海量数据的存储和检索，提高数据处理效率。智库研究大数据可视化算法基于经典算法改进优化得出，如分类、决策树、朴素贝叶斯、Apriori 算法、神经网络、遗传算法等。主要通过在经典算法的基础上在计算机处理技术和图形分析技术上做一定的改进，对大数据可视化的智库研究起到促进作用。如目前常用的机器学习算法，通过深度学习，可预测基于海量数据的不确定事件的发展趋势，图形表达内容丰富，用户交互性能良好 [25]。

复杂智库信息的图形可视化技术根据数据类型不同呈现的可视化结果也不同，其涉及的数据包括文本数据、网络数据、时空数据、多维数据等，并具有实时交互的功能 [26-27]。目前，虚拟现实技术、

可视化仿真等技术都在智库复杂信息可视化中有所运用，表现内容更加全面，交互形式更加丰富，促进大数据可视化的智库研究快速发展。大数据环境下，支持智库可视分析的人机交互技术逐渐成为大数据可视化的智库研究中的重要方法 [28]。

随着智库研究中处理信息量的迅速增加以及资源共享的需求，可视化的智库研究中分布式处理技术逐渐被运用。分布式计算如云计算技术，其深度学习功能能进行图像识别、语音识别、数据分析等 [29-30]。且具有快速访问资源和多用户共同使用的特点，且分布范围广，使用人数多，网络系统性能强，广泛运用于目前大数据可视化的智库研究中，有助于具体化和多维化共享数据可视化分析结果。

3 大数据可视化视角下的智库应用研究分析

目前世界各国越来越重视智库建设，将智库研究成果视为政治、经济决策的重要依据。随着大数据时代的到来，信息量不断激增，如何有效的进行海量数据采集，并进行有效、动态、直观的分析成为智库研究的挑战。大数据可视化的智库研究方法是目前智库研究中比较新兴的方法。目前，社会网络分析、知识图谱、共现网络分析等都是目前智库研究中较为常用的可视化方法，不仅对智库研究起到了强有力的促进作用，且有利于数据的高效和精准采集、处理、分析、预测。因此，文章对大数据可视化视角下智库应用研究和创新进行分析，并用 CiteSpace II 软件对知网里的部分智库研究期刊文章展开可视化分析和探索。从智库数据的快速高效处理、智库研究方法的决策科学准确化、智库报告中复杂图形可视化应用、智库研究成果的复杂问题简单化、支持智库研究过程中可视分析的用户交互、智库数据资源集聚共享的方法创新、大数据可视化提升智库作用和影响力等 7 方面来探讨大数据可视化视角下的智库应用研究和创新。

3.1 智库数据的快速高效处理

大数据可视化的智库研究方法在数据处理阶段较传统的方法具有更快的速度和更高的效率，其采集的智库数据包括传统的结构化数据和海量非结构化数据，较传统的统计数据全面而具体。

文章的数据来源于 CNKI 中的 CSSCI 数据库，截止时间为 2017 年 7 月 19 日，精确检索到 2000 年~2017 年篇名中含有“智库”的期刊文章共 795 篇。通过将数据导入到 CiteSpace II 中进行数据转换、处理，使之快速成为可视化处理的标准格式数据，选择“keyword”选项，设 selection criteria 参数为 top50，Threshold interpolation 为（2，1，10；2，1，10；2，1，10），选择 Pruning sliced networks 和 Minimum spanning tree 方法对关键词字段识别和可视化分析，并快速得到了分析结果。对文献数据的可视化分析的空间状态用数据形式表示如表 1 所示。

表 1 文献数据可视化分析空间状态数据表示

Table1 The data representation of the spatial state of the literature data visualization analysis

时间段	阈值调谐数	年段内引文数	年段内聚类节点	年段内连线数
2007	2 | 1 | 0.1	5	2	1/1
2008	2 | 1 | 0.1	16	0	0/0
2009	2 | 1 | 0.1	83	5	2/2
2010	2 | 1 | 0.1	136	10	12/12
2011	2 | 1 | 0.1	69	5	4/4
2012	2 | 1 | 0.1	229	27	43/43
2013	2 | 1 | 0.1	222	16	15/15
2014	2 | 1 | 0.1	425	51	102/113
2015	2 | 1 | 0.1	573	77	154/244
2016	2 | 1 | 0.1	736	95	190/308
2017	2 | 1 | 0.1	281	23	27/27

除 2001 年段内引文数为 10，2000 年~2006 年的其余数据均为 0，因此表中只列出了 2007 年~2017 年的数据。从表 1 中可以明显看出，采集到的智库研究相关文献引文数从 2008 年至 2017 年均呈递增的态势，其中 2016 年的引文数最高为 736，说明 2007 年以后，国内智库研究的热度一直在增加。聚类节点和年段内连线数的逐年递增也可间接说明 CSSCI 数据库中关于智库研究的文章数量在逐渐增加，反

映该研究趋势逐渐增强。大数据可视化方法在传统智库研究方法基础上，可快速处理分析大量智库数据，并及时得出客观详细的数据分析结果，对高效智库研究具有一定支持作用。

3.3 智库报告中复杂图形可视化应用

大数据可视化方法在智库研究中的运用弥补了传统方法在图形可视方面的不足，其运用计算机技术将大规模集群的信息有效处理，将数据之间的相关关系运用多维可视图形的形式展现出来，尤其是智库研究报告中的复杂图形可视化应用，使得用户可以直观的看到数据之间的关系以及数据与数据之间构成的整体性网络。目前，虚拟现实技术、可视化仿真等技术都在快速发展，在智库报告中逐渐有所展现，表现内容更加全面，交互形式更加丰富。

世界资源研究所（WRI）在其自建的数据库中运用大数据可视化技术帮助用户快速查询研究领域分布地图、研究热点图、时空图等可视化图形，有助于用户在查询相关报告时直观准确地观察研究领域的地域分布、热点、趋势等信息。中国知网（CNKI）在其数据库和相关报告中也运用大数据可视化方法进行文献可视化分析，以复杂图形展示报告内容，帮助用户了解和认识引文网络、关联作者、相似文献等可视化图形，直观展示研究热点、相关研究人员、延伸文献等信息。目前国内外智库研究报告中，科学知识图谱是较为常用的可视化分析方法，重要的节点和节点间关联关系能通过图形直观显示，有助于把握智库研究趋势。通过将采集的文献信息运用 CiteSpace II 中共现网络的形式展现出来，用户可直观的看到整个网络系统的关键词，通过所显示的节点大小从而直观判断智库研究中的热点问题和趋势，可视化结果如图 2 所示。

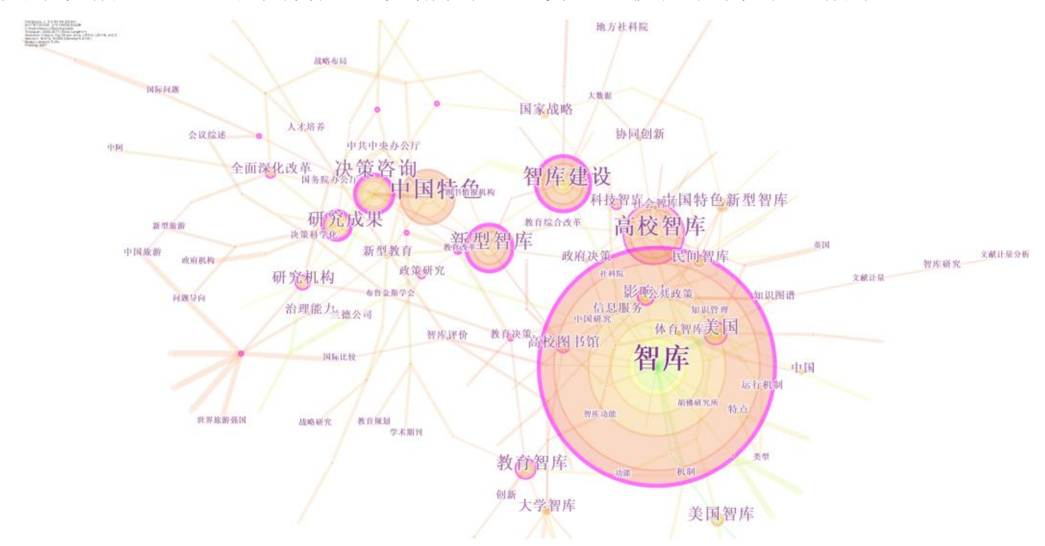


图 2 关键词共现网络图

Figure 2 The co-occurrence network graph of keywords

图 2 中的节点大小表示与智库研究有关的关键词的频数大小，节点之间的连线表示关联度，2000 年以来智库研究文献中出现频率最高的关键词依次有“智库”“高校智库”“中国特色”“智库建设”“新型智库”“决策咨询”“研究成果”等，从节点大小判断这些关键词均是目前智库研究的热点。在输出图形同时还附带了相关的具体图形数据，包括具体的关键词、频数、中心度、年份等，如表 2 所示。

表 2 关键词频数和中心度

Table 2 The frequency and centrality of keywords

频数	中心度	年份	关键词
200	0.37	2007	智库
52	0.17	2014	高校智库
50	0.07	2013	中国特色
46	0.25	2012	智库建设
38	0.20	2014	新型智库

34	0.35	2012	决策咨询
24	0.29	2012	研究成果
23	0.17	2007	美国
19	0.11	2014	教育智库
15	0.03	2010	美国智库
13	0.04	2014	中国特色新型智库
13	0.14	2012	研究机构
13	0.13	2014	影响力
12	0.08	2014	大学智库
12	0.07	2013	民间智库
10	0.18	2015	高校图书馆

表 2 中按关键词出现的频数递减的顺序排序判断近几年国内智库研究的热点，关键词“智库”的频数为 200，中心度为 0.37，结合图形中焦点颜色的变化发现，含该关键词的文献时间跨度最长，而其余的除了“决策咨询”的中心度为 0.35 之外，其余关键词的频数和中心度逐渐减小，时间跨度缩短。直观说明国内关于智库的研究从 2007 年开始逐渐增多，并衍生出多个研究分支方向。复杂图形可视化直观展示相关文献关键词的内容、分布、关联情况及节点特征，从焦点圆心到圆周的年轮颜色变化可判断历年智库研究的热点和趋势。复杂图形可视化虽然需要依靠大量数据支持，但其图形化信息的直观展示在智库研究报告中的应用有助于公众和研究人员具体化理解抽象问题，提高研究效率。

3.2 智库研究方法的决策科学准确化

大数据可视化的智库研究方法依据的是科学的研究方法，在智库数据的统计基础上，依托计算机技术和大数据可视化算法实现海量智库数据的高效、科学分析，保证决策科学准确化。在对智库研究文献的关键词共现网络分析的基础上，对分析结果进行聚类算法分析，并展示可视化结果，如图 3 所示：

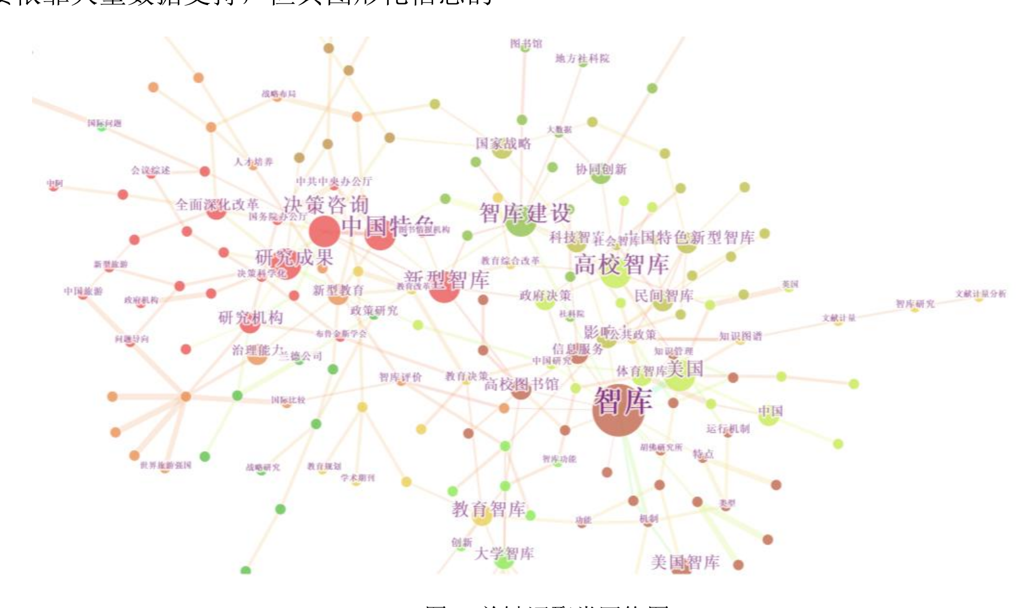


图 3 关键词聚类网络图

Figure 3The cluster network graph of keywords

从图 3 中可以看出，通过聚类算法将关键词的类别用不同颜色显示在可视化图形中，可直观的看出在检索到与智库研究有关的文章中，文章的高频关键词大致可分为智库、智库建设和高校智库、中国特色和决策咨询三个类别，也就是说不通颜色的节点区域表示智库研究的侧重点不同。因此可以在判断智库研究热点的基础上，根据聚类算法得出的可视化结果能看到智库研究的主要分支方向和侧重点。科学的计算方法有助于快速、高效、准确的帮助智库研究人员进行科学分析。2015 年 1 月 19 日，

上海交通大学电子信息与电气工程学院大数据工程技术研究中心成立，其目的之一就是在大数据环境下，努力建设大数据产业顶尖智库，在加强大数据产业快速发展的同时，增强智库建设，提升智库的社会影响力。上海交大的科研人员根据大数据分析算法，从高维空间分析数据，根据数据本身的规律以及数据之间的结构和关系，发掘最本质的信息[31]。因此，数据处理流程的严格设置以及大数据可视化算法的优化，有助于科学地完善智库研究流程，可视化展示分析结果，显示主要观点，提高决策的科学性和准确性。

3.4 智库研究成果的复杂问题简单化

传统的分析方法能对数据进行抽象和简化，但是却很难准确表示出智库数据之间的关系和结构，对智库成果的描述也具有专业性，不易理解。而大数据可视化方法就是让用户从直观展示的图形信息中高效获取知识资源，能用多维图形展现智库研究成果，如层次与网络结构可视化、时空可视化等具体图形信息，有助于简化复杂问题，高效理解分析。

以大数据可视化为依托的智库研究可有效的展示智库研究成果，将众多专业性较强的结果运用图形的形式直观展示，大大简化了问题，便于更多的用户进行准确理解和宏观分析，从而共同参与政策制定和智库研究。国际主要智库每年都会发布与智库研究有关的研究报告，大数据可视化方法则被运用到报告中将研究成果可视化展示，简化复杂问题，便于公众阅读。英国皇家国际事务研究所在其发布的研究报告中就运用知识地图的方式直观、简单、高效地展示研究成果。近年来大数据可视化软件的开发和应用为国内外智库研究提供了便捷，大量复杂的数据可通过软件进行运算，生成图形直观展示计算结果，简化分析流程，提高决策效率。

由此看来，大数据可视化的智库研究方法能将智库研究成果中复杂的问题以简单易懂的方式呈现出来，使得大众对智库研究的结果更容易接受，从而增加大众对政府政策制定过程的参与和对智库研究成果的判断和认可。以时序可视化分析为例，其在智库研究中也较为常用，可以此判断各个年份的研究态势，该可视化结果如图 4 所示。

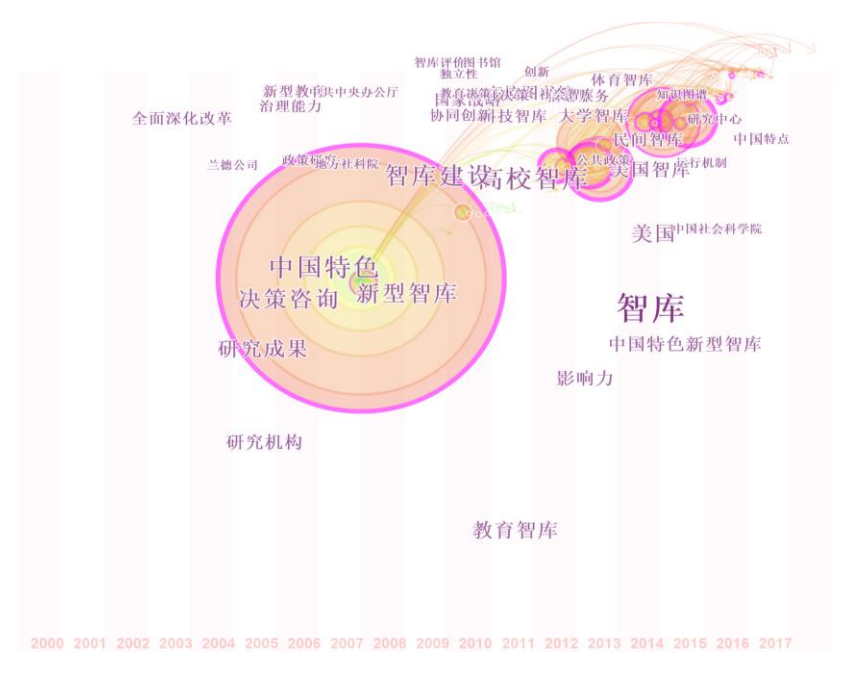


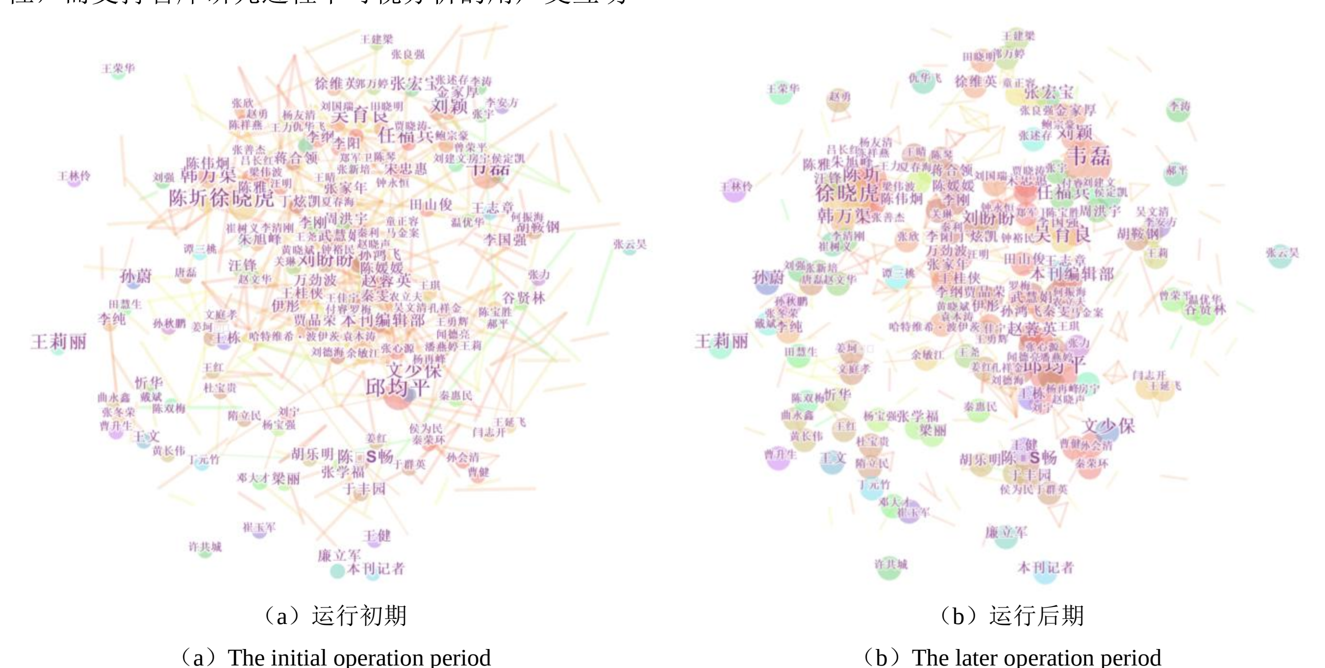
图 4 关键词时序可视化分析

Figure 4 The visualization analysis of time series of keywords

从图 4 中可清晰看出，2000 年开始智库研究的热点逐渐从决策咨询、新型智库到智库建设，2010 年后逐渐从高校智库向中国特色新型智库过渡。说明我国的智库研究在过去几年的发展下，有了多个新的研究分支，发展态势良好。不同于传统的文献计量统计的繁琐、复杂，时序可视化方法通过设定一定时间区间将复杂的问题直观的展示出来，完整的表现出了抽象的数据时序变化情况，乃至具体细节，包括焦点大小，关联关系等，丰富了信息含义，简化了分析过程，直观展现目前我国智库研究的发展趋势，使得结果更容易为大众接受，是大数据可视化方法在智库研究中的应用创新之一。

3.5 支持智库研究过程中可视分析的用户交互

智库研究的趋势和热点能被大数据可视化展示，但为了让用户更易于理解和分析，用户交互功能被运用到大数据可视化的智库研究中，可对数据或者图形结果进行一定的处理和操纵来分析和理解数据。现如今海量的非结构化数据使得传统的可视化分析方法在用户理解数据方面存在一定的局限性，而支持智库研究过程中可视分析的用户交互功能可改善这种情况。科学知识图谱软件在智库研究的可视分析中提供良好的用户交互功能，是大数据可视化的智库研究方法之一。通过对智库研究的文献作者在 CitespaceⅡ软件中进行可视化分析，生成文献作者共现网络，并截取网络图的动态运行过程图，根据需求实时改变节点大小突出焦点，如图 5 所示。



（a）运行初期

（a）The initial operation period

（b）运行后期

（b）The later operation period

图 5 文献作者共现网络

Figure 5 The co-occurrence network graph of literature authors

文献作者共现网络由交互界面动态展示，图 a 为网络图运行初期，节点之间的关系连线比较混乱，焦点不明显。图 b 为网络图运行后期，节点之间的关系连线逐渐清晰，且通过文字大小和节点大小的调节，显示了主要作者分布，突出了网络焦点。可视化的交互界面有助于提高智库研究的灵活性和高效性，信息的关系图谱能通过交互的功能将数据信息传递给用户。

在 2017 年 1 月 7 日举办的第二届新经济智库大会上，卡内基梅隆大学计算机学院副院长 Justine Cassel 认为，下一代人机交互界面需具备两个属性：通过对话和肢体语言进行自然互动；采用社交模式而非任务模式 [32]。直观、自然的人机交互有助于未来智库平台完善和技术创新。兰德公司（RAND）拥有美国多个行政机构及行业的数据，信息需求者、供应者、多个数据库之间的相互良好交互有利于增强后续分析的灵活性。大数据可视化交互功能使得智库研究的分析过程更加灵活、直观、便捷，同时还保证了结果的准确性和易接受性。

3.6 智库数据资源集聚共享的方法创新

智库数据资源集聚共享是未来智库发展的必然趋势。智库研究通常涉及访问众多的知识型数据库，从数据库中获取大量信息并进行处理具有一定的难度，由于机构之间的数据相对独立，资源共享有限。并且公众用户用传统的查询、统计等方法无法直观的判断智库研究机构的研究现状、机构之间的联系以及数据库资源等信息。随着互联网和多媒体技术的应用和普及，各智库研究机构纷纷运用互联网平台进行数据采集、分析、可视化成果展示，并进行智库数据资源共享，实现机构间互通有无，普通大众共同参与学习和研究。

目前，很多智库都运用分布式处理技术建立知识型数据库，并整合相关的数据资源，建立开放式的数据资源平台，运用大数据可视化分析方法直观展示数据资源分布，使大众参与分析和判断，高效

获取智库数据资源，实现智库数据资源共享。很多国际知名智库都建立了强大的数据库，以便保存重要的文献资料和研究报告，同时提供网络资源共享平台，加强知识的学习和共享。英国皇家国际事务研究所在完善建设数据库的基础上，设立专门的数字资源部门支持资源共享。卡内基国际和平基金也提供智库知识资源的共享服务，服务对象是其注册会员。中国社会科学院的中国社会科学网数据中心就将社会科学领域的学术期刊、社会调查资料等整合、集聚在一个开放平台中，实现资源共享。同时，云平台的建立和相关配套系统的开发应用使智库的存储和利用效率大大提升，数据库之间的信息通信更加便捷，能使用户便于访问数据库，降低了信息的获取成本，有效实现资源共享，满足用户需求。

在大数据可视化软件的支持下，用户可便捷识别智库研究相关机构，从而了解我国智库研究机构的研究现状、机构之间的联系，以此间接判断机构之间的合作和资源的共享情况。从图 6 中可以看出，在所选文献的范围内，对智库做了相关研究较多的研究机构主要有中国科学院大学、中国科学院文献情报中心、南京大学信息管理学院、武汉大学信息管理学院、武汉大学信息资源研究中心、清华大学公共管理学院、北京大学信息管理系等。图中节点颜色不同，说明节点类别不同，少数节点之间有表示关联关系的连线。在智库研究机构可视化结果中，我国智库研究机构数量可观，但机构之间的联系较弱，机构之间存在一定的独立性。因此为了我国智库的深入研究和建设，机构之间应加强合作，共享资源，避免孤岛效应。应使用户不仅能便捷判断并访问主流的研究机构和数据库，而且还能根据大数据的云服务功能更好地共享智库数据和可视化分析结果，实现资源的无缝沟通，更加高效、准确地进行资源共享。



图 6 智库研究发文机构共现网络

Figure 6 The co-occurrence network of institutes that published articles in think tanks research

3.7 大数据可视化提升智库作用和影响力

宾夕法尼亚大学智库与公民社会项目主任詹姆斯・麦甘博士曾表示，大数据不仅可以用来收集海量数据，但更重要的是要对海量数据进行专业分析通过数据分析结果发掘社会问题，如健康、交通等。为政府和研究人员提供未来处理问题的决策依据。大数据技术的良好利用能助推智库的作用和影响力提升。

大数据可视化分析方法在智库研究中应用创新，使得智库研究人员可以充分将互联网、计算机技术、数据库技术等相结合，进行深入的智库研究，

运用大数据可视化分析方法分析问题，并帮助决策人员和公众理解问题、获取知识资源，助推政府在解决社会问题上科学决策。近年来、大数据可视化领域逐渐衍生出大量的新兴技术，这对未来智库研究具有很好的促进作用。但在社会层面上，大数据可视化分析技术需要有针对社会公众的人性化的交互系统，不仅使得智库工作者快速掌握大数据可视化分析技术，满足智库研究人员的需求，同时使更多的社会公众理解问题、参与政府决策，促进公众认知和学习智库研究成果，实现智库数据资源共享，有力提升智库的社会影响力。

4 结论

文章对智库研究和大数据可视化的研究现状进行阐述，整理大数据可视化的智库研究方法相关理论，并分析在大数据可视化分析支撑的智库研究的应用和创新，结果表明：

（1）智库建设逐渐受到各国政府的广泛重视，越来越多的学者针对智库展开相关研究，智库研究逐渐成为学术界研究的热点。但我国的智库研究方法还比较单一，未形成完整的研究体系。大数据时代下海量信息的收集、处理、分析成为难点，大数据可视化分析方法在传统可视化方法的基础上可高效处理海量数据、在图形、空间直观展示数据分析结果，具有一定的结果可接受性，有必要将大数据可视化方法运用到智库研究中。

（2）智库的研究方法主要基于传统的计量分析方法演变而来，但面对智库海量信息，大数据可视化的智库研究方法快速发展。目前大数据可视化的智库研究方法主要包括智库海量非结构化数据处理技术、智库研究大数据可视化算法技术、复杂智库信息的图形可视化技术、支持智库研究可视分析的人机交互技术、可视化的智库研究中分布式处理技术等。

（3）大数据可视化技术目前都被逐渐运用在智库研究中，给智库研究的方法和技术创新提供了强有力的支撑。海量智库数据可以通过计算机软件平台快速准确地获取，并进行高效的数据预处理。科学的数据采集、处理、分析的方法使智库研究方法决策科学准确化。运用多维的图形界面将复杂的数据结构通过可视化形式展现出来，使复杂问题简单化。同时良好的交互技术使得用户能更加准确的操纵、分析、理解数据。基于大数据的云共享技术能帮助更多的用户访问知识型数据库，实时采集数据，分析结果，有效实现资源共享。

（4）大数据可视化支撑的智库研究方法目前已有一定的应用，是传统研究方法的重大改进，对智库研究的应用和创新都具有强有力的支撑作用，这不仅是当前智库研究方法的初探，也是该项研究未来的发展趋势。

参考文献：

[1] 许鑫，吴珊燕. 智库知识库的构建研究[J]. 情报理论与实践，2014，3：68-72.

[2] 中共中央办公厅，国务院办公厅. 关于加强中国特色新型智库建设的意见[EB/OL]. [2017-06-10]. http://www.gov.cn/xinwen/2015-01/20/content_2807126.htm.

[3] Mcgann J G. 2015 Global Go To Think Tank Index Report[EB/OL]. [2017-05-18]. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1009&context=think_tanks.

[4] Peter J. Jacques, Riley E. Dunlap, Mark Freeman. The Ordanization of Denial: Conservative Think Tanks and Environment Scepticism[J]. Envoronmental Politics, 2008, 17(3):349-385.

[5] 吕红. 国际智库研究态势及其热点主题分析[J]. 高教发展与评估，2016, (05):42-49, 83, 121.

[6] Gershon N D, Eick S G. Information Visualization[J]. IEEE Computer Graphics and Applications, 1997, 17(4):29-31.

[7]Peck R, Olsen C, Devore J. Introduction to Statistics and Data Analysis(3 edition)[M]. Australia, Belmont, CA, 2008.

[8]Lifschitz V, Harmelen F V, Porter B. Handbook of Knowledge Representation[M]. Elsevier, 2008, 2(4):374–391.

[9]Frey S, Sadlo F, Ma K L, et al. Supporting Communication and Coordination in Collaborative Sensemaking[J]. IEEE Transactions on Visualization and Computer Graphics. 2014, 20 (12):1633-1642.

[10]Leung Y K, Apperley M D. A Review and Taxonomy of Distortion-oriented Presentation Techniques[J]. ACM Transactions on Computer-Human Interaction. 1994, 1(2): 126-160.

[11]Drigas A, Koukianakis L, Papagerasimou Y. Towards an ICT-based Psychology: W-psychology[J]. Computers Human Behavior. 2011,27:1416– 1423.

[12]Boukhelifa N, Rodgers P J. A Model and Software System for Coordinated and Multiple Views in Exploratory Visualization [J]. Information Visualization, 2003,2(4):258–269.

[13]俞宏峰. 大规模科学可视化[J]. 中国计算机学会通讯，

2012, 8(9): 29-36.
[14] 陈 明 . 大 数 据 可 视 化 分 析 [J]. 计 算 机 教 育,2015,(05):94-97.
[15]Agrawal R, Srikant R. Fast Algorithms for Mining Association Rules in Large Databases[C]. Proceedings of the 20th International Conference on Very Large Data Bases, VLDB'94, Santiago, Chile, 1994.
[16]Dilla W, Janvrin D J, Raschke R. Interactive Data Visualization: New Directions for Accounting Information Systems Research[J]. Journal of Information Systems American Accounting Association. 2010, 24(2):1-37
[17]Roger S. Software Engineering: A Practitioner's Approach [M]. New York: McGraw Hill. 2010.
[18]Card S K, Mackinlay J D. Shneiderman B. Readings in information visualization: using vision to think[M]. Readings in information visualization: Morgan Kaufmann Publishers, 1999: 647-650.
[19]Abelson D E. American Think-Tank and Their Role in US Foreign Policy[M]. New York: ST. Martin's Press, 1996.
[20]Rich A. Think Tanks, Public Policy, and the Politics of Expertis[M]. New Yorks: Cambirdge University Press, 2004.
[21]赵蓉英,魏绪秋. 基于文献数据的国内外智库研究动态分析[J]. 智库理论与实践,2016,(03):6-14.
[22]吕青. 从智库研究理论到科技智库建设：专访中国科学院科技战略咨询研究院潘教峰院长[J]. 智库理论与实践,2016,(06):2-5.
[23]武慧娟,秦雯,韩林丛,孙鸿飞. 国内外智库研究态势知识图谱对比分析[J]. 现代情报,2016,(11):170-177.
[24]陈臣. 基于Hadoop的图书馆非结构化大数据分析与决策系统研究[J]. 情报科学,2017,(01):24-28.
[25]何清,李宁,罗文娟,史忠植. 大数据下的机器学习算法综述[J]. 模式识别与人工智能,2014,(04): 327-336.
[26]Viégas F B, Wattenberg M, Feinberg J. Participatory Visualization with Wordle[J]. IEEE Transactions on Visualization & Computer Graphics, 2009, 15(6): 1137-1144.
[27]Zhao J, Chevalier F, Collins C, et al. Facilitating Discourse Analysis with Interactive Visualization[J]. IEEE Transactions on Visualization & Computer Graphics, 2012, 18(12): 2639-48.
[28]Oviatt S, Cohen P, Wu L, et al. Designing the User Interface for Multimodal Speech and Pen-Based Gesture Applications: State-Of-The-Art Systems and Future Research Directions [J]. Human-computer Interaction, 2000, 15 (4): 263-322.
[29]Chan T H, Jia K, Gao S, Liu J, et al. PCANet: A Simple Deep Learning Baseline for Image Classification[J]. IEEE Transactions on Image Processing, 2014,24 (12):5017－5032.
[30] Hinton G, Deng L, Yu D, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups[J]. IEEE Signal Process Magazine. 2012,29 (6): 82－97.
[31]中国新闻网. 上海交大打造产业顶尖智库 推动大数据产业发展[EB/OL]. [2017-05-19].http://www.chinanews.com/edu/2015/01-19/6983566.shtml.
[32]国际在线. 第二届新经济智库大会召开人机交互、平台经 济 体 受 热 [EB/OL]. [2017-06-07]. http://news.cri.cn/20170107/95f2415b-7777-78e8-ff4b-4f9e9435b60e.html.

作者贡献说明：

熊励：负责研究选题、研究与方法设计及论文撰写；
王锟：资料收集与论文撰写；
赵明辉：文献综述与论文撰写。

收稿日期：2017-06-21　修改日期：2017-07-20　本文责任编辑：唐果媛